RAHUL PODUGU

Software Engineer (Backend/Distributed Systems)
443-983-5865 | rahulpodugu2@gmail.com | linkedin.com/in/rahulpodugu | github.com/Rahul2251999

Summary

Software Engineer with over 2 years of experience designing and building scalable backend systems and distributed architectures. Proficient in Java and Spring Boot, with strong skills in system design, algorithms, and concurrency. Experienced in developing AI-driven platforms and data-intensive financial applications that support real-time analytics, automation, and intelligent decision-making. Passionate about delivering secure, high-performance solutions that advance innovation in FinTech and AI-powered software ecosystems.

Technical Skills

Programming Languages: Java, Python, C++, C#; JavaScript/TypeScript, HTML, CSS

Frameworks & Technologies: Spring Boot, Hibernate; RESTful APIs, gRPC, GraphQL; Apache Kafka, Apache Spark, Redis, React

Cloud & DevOps: AWS (EC2, S3, Lambda), Docker, Kubernetes; CI/CD (Jenkins, GitLab CI)

Databases & Big Data: MySQL, MongoDB, Cassandra; Hive, Presto, Spark SQL

Systems & Tools: Microservices Architecture, Distributed Systems; Linux, Git, Agile/Scrum; Concurrency & Multithreading; Networking protocols (TCP/IP, HTTP); Observability/Monitoring (Prometheus, Grafana, ELK)

Professional Experience

Aztra – Columbus, OH – Software Engineer II

May 2024 – Present

- Engineered and deployed **10+ scalable microservices** with Java Spring Boot and gRPC on Kubernetes, enabling the AI analytics platform to handle **100K+ events/sec** and boosting throughput by ∼30%.
- Optimized Kafka + Spark–based real-time data pipeline processing **500GB+/day**, cutting processing latency by 40% and accelerating delivery of client-facing analytics.
- Built custom **Prometheus metrics exporters** and Grafana dashboards with automated alerting, reducing mean time to resolution (MTTR) by 50% and improving uptime to 99.9%.
- Integrated Redis-based caching layer into backend microservices, reducing **p95 API latency by 35%** and cutting database load by 40%.
- Led on-call rotations and performed deep-dive root cause analyses of **10+ Sev-1 incidents**, driving permanent fixes that reduced repeat issues by 60%.
- Partnered with cross-functional teams in bi-weekly Agile sprints, leveraging CI/CD pipelines and TDD to deliver **3 major features per quarter** while maintaining high code reliability.

University of Maryland, Baltimore County (UMBC) – Baltimore, MD – Software Engineer / Research Assistant

Sep 2022 – Nov 2023

- Built a distributed simulation platform in Java/Python to evaluate load-balancing algorithms in CDN research, enabling real-time analysis of system behavior under **10K+ concurrent requests**, node failures, and traffic surges.
- Designed predictive models in Python and Spark to forecast content popularity, improving cache pre-fetching efficiency by **20%** and reducing content retrieval latency.
- Implemented a fault-tolerant, low-latency video streaming prototype on AWS using EC2 and S3, resilient to **single-node failures** through automated recovery mechanisms.
- Co-authored a peer-reviewed research paper on distributed systems, introducing novel load-balancing algorithms that increased system throughput by **15%** while enhancing fault tolerance; presented results at departmental seminars.

UST – Bangalore, India – Software Development Engineer

Aug 2021 – Jun 2022

- Developed and maintained backend services for a fintech client using Java Spring Boot and MySQL, designing RESTful APIs that processed **1M+ daily transactions** with strong consistency guarantees.
- Optimized database performance by profiling and refactoring slow queries, implementing Redis caching strategies that boosted application throughput by **25%** and reduced page load times.
- Designed and documented REST APIs in Spring Boot to integrate with React.js frontends, enabling real-time transaction dashboards and improving client response times by **30%**.
- Built new microservice modules and enhanced inter-service communication via gRPC, writing unit and integration tests that reduced production defects by **40%**.
- Refactored monolithic components into AWS-hosted microservices, improving scalability and maintainability; authored detailed migration documentation to guide future transitions.

Fidelity National Financial – Bangalore, India – Software Engineer Intern

Apr 2021 – Aug 2021

- Created a Java Spring Boot proof-of-concept module to automate insurance claim document processing, reducing manual effort by **30%** and accelerating claims resolution.
- Enhanced an internal claims system by refactoring backend logic and optimizing SQL queries, cutting page load time by **20%** and improving responsiveness.
- Built responsive React.js components with dynamic form handling and real-time validation, digitizing workflows and reducing manual data entry errors by **25%**.
- Supported DevOps automation by designing CI/CD pipelines with Jenkins and containerizing applications via Docker, reducing deployment time from **hours to minutes**.
- Collaborated in Agile sprints with senior engineers through code reviews and pair programming, contributing bug fixes and enhancements that improved overall code quality.

Projects

SecureShare – Decentralized P2P File Sharing System

Developed a decentralized peer-to-peer file sharing system that combined **RSA and AES encryption** to ensure secure data transfer across distributed nodes. Designed chunk-based distribution and a **Distributed Hash Table (DHT)** for fault-tolerant peer discovery and efficient file lookup, achieving **3**$\times$ **faster retrieval speeds** compared to baseline P2P implementations while maintaining strong encryption standards.

ChatAssist – AI-Powered Conversational Agent

Created an AI-powered chatbot using **LangChain and OpenAI APIs** to provide real-time incident resolution support for on-call engineers. Integrated a **Milvus vector database** for semantic search across runbooks and root-cause analyses, with keyword-based fallback for low-confidence matches. This reduced average incident resolution time by **35%** by generating contextual playbooks with validation steps, rollback procedures, and citations.

StockStream – Distributed Real-Time Stock Ticker

Designed a distributed, event-driven stock ticker system capable of streaming live price updates to **thousands of concurrent clients**. Leveraged **Apache Kafka** for ingestion and **Redis** for caching, enabling sub-100ms propagation of updates across subscribers. Optimized consumer partitioning and indexing improved throughput by **2.5**$\times$ and reduced p95 latency from 800ms to 250ms.

Education

Master of Science in Computer Science, University of Maryland, Baltimore County – Graduated 2024
Coursework in Distributed Systems, Cloud Computing, Algorithms, Big Data Analytics.